

# Special Relativity Derived from Spacetime Magma

Fred Greensite\*

Department of Radiological Sciences, University of California Irvine, Irvine, California, United States of America



## Abstract

We present a derivation of relativistic spacetime largely untethered from specific physical considerations, in contrast to the many physically-based derivations that have appeared in the last few decades. The argument proceeds from the inherent magma (groupoid) existing on the union of spacetime frame components  $\mathbb{R}$  and Euclidean  $\mathbb{R}^3$  which is consistent with an “inversion symmetry” constraint from which the Minkowski norm results. In this context, the latter is also characterized as one member of a class of “inverse norms” which play major roles with respect to various unital  $*$ -algebras more generally.

**Citation:** Greensite F (2014) Special Relativity Derived from Spacetime Magma. PLoS ONE 9(6): e100583. doi:10.1371/journal.pone.0100583

**Editor:** Gerardo Adesso, University of Nottingham, United Kingdom

**Received:** February 6, 2014; **Accepted:** May 28, 2014; **Published:** June 24, 2014

**Copyright:** © 2014 Fred Greensite. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** The author has no support or funding to report.

**Competing Interests:** The author has declared that no competing interests exist.

\* Email: fredg@uci.edu

## Introduction

In an effort to acquire a deeper understanding of its necessity, in recent years there have been a number of alternative derivations of relativistic spacetime using differing combinations of physical principles distinct from those employed in the original electro-dynamical presentation. These have variously included arguments related to the consistency of Newton’s First Law and spacetime isotropy with non-instantaneous interaction propagation [1], a probability distribution resulting from quantum fluctuations of the spacetime metric [2], presence or absence of practicable prediction based on the relevant resulting well-posed versus ill-posed field equations [3], requirements for the existence of observed elementary particles and dynamics in the context of basic quantum mechanical considerations [4], implications of electric/magnetic reciprocity [5], and the effective ill-posedness of the Einstein field equations unless the Minkowski metric pertains locally [6] (with the above citations representing only a sample of the relevant work published on this topic). We here augment these efforts by presenting a derivation unencumbered by such physical principles, relying instead on more primitive mathematical considerations - in particular, a symmetry deriving from the inherent magma (or groupoid) implied by a spacetime frame. Perhaps this can be seen as support for recent sentiments that reality accessible to physical investigation is little more than mathematical structure [7,8]. It also allows alternative placement of relativistic spacetime as prior to, rather than deriving from, the various physical formalisms. A further consequence of our presentation (developed in some detail) is the attendant understanding of the Minkowski norm as one member of a class of “inverse norms”, which are important features of various unital  $*$ -algebras more generally.

## Results

### A magma-derived spacetime algebra

A magma is simply a set with a product that maps any pair of elements of the set into the set [9].

**Derivation of the Minkowski metric from the magma of space and time.** It would seem pointless to pursue physics in

the absence of an assumption confirming the possibility of a local “observer” within a space, capable of making a series of “physical predictions” and executing a series of subsequent measurements evaluating those predictions. Notwithstanding speculations concerning a universe with multiple time dimensions, and (indeed) timeless formulations [10], we will interpret the prior sentence as indicating an observer-dependent co-dimension-1 foliation of spacetime with the leaves parameterized by “time” (this is formalized as an axiom in [5]). More specifically, we assume that “phenomena” in affine spacetime can be described in the context of any embedded coordinate frame that indexes the points as  $\mathbb{R} \oplus \mathbb{R}^3$  where  $\mathbb{R}^3$  is Euclidean. This is in line with the latter’s role as the global coordinate system underlying all established (i.e., experimentally supported) nongravitational physics, and as the local coordinate system in general relativity (i.e.,  $\mathbb{R} \oplus \mathbb{R}^3$  is the coordinate system of the tangent space at each point of the classical spacetime pseudo-Riemannian manifold). Since  $\mathbb{R}$  and  $\mathbb{R}^3$  are  $\mathbb{R}$ -vector spaces, and  $\mathbb{R}^3$  is endowed with the Euclidean inner product, it follows that  $\mathbb{R} \cup \mathbb{R}^3$  is a magma, i.e., the magma product  $\star$  is given by the three preexisting vector space product operations, so that for  $\tau, \tau_1, \tau_2 \in \mathbb{R}$ , and  $s, s_1, s_2 \in \mathbb{R}^3$  we have  $\tau_1 \star \tau_2 \equiv \tau_1 \tau_2 \in \mathbb{R}$ ,  $\tau \star s \equiv \tau s \in \mathbb{R}^3$ , and  $s_1 \star s_2 \equiv s_1 \cdot s_2 \in \mathbb{R}$ . Thus, from their associations with  $\mathbb{R}$  and Euclidean  $\mathbb{R}^3$ , time and space are already thoroughly linked (or jumbled together) prior to taking their direct sum. Via the obvious vector space isomorphisms, the spacetime points comprising the union of the two spacetime frame components  $\mathbb{R} \oplus \{0\}$ ,  $\{0\} \oplus \mathbb{R}^3$ , is then also a magma, with  $(\tau_1, 0) \star (\tau_2, 0) = (\tau_1 \tau_2, 0)$ ,  $(\tau, 0) \star (0, s) = (0, \tau s)$ , and  $(0, s_1) \star (0, s_2) = (0, s_1 \cdot s_2)$ . The magma exists for any spacetime frame of the specified form, and for any pair of spacetime points there are such frames with respect to which their magma product can be calculated. Elevating the magma products to the status of phenomena (because the quotient, as the operation inverse to the product, is intrinsic to description of the physical phenomenon of velocity), the product of any two spacetime points should be able to be calculated in some sense in the context of any frame (in accordance with the assumption made earlier in this paragraph;

see also the subsection following the next one). This would be equivalent to associating each spacetime point with a map of spacetime to itself. It is significant that a linear map is defined by the product of any fixed point of the magma with the points of either of the magma's linear subsets ( $\mathbb{R} \oplus \{0\}$  or  $\{0\} \oplus \mathbb{R}^3$ ). There is a unique elaboration of the product operation such that this linearity property pertains over all spacetime points in the context of the frame - i.e., so that the magma product generalizes in a manner consistent with the linear properties of the space  $\mathbb{R} \oplus \mathbb{R}^3$ , and so that spacetime nonlinearities do not exist (e.g., consistent with the prior notion that the spacetime frame be 'flat'). From this linearity requirement, the product of any two points with respect to a frame of the above type conforms to the distributive law, and is given by

$$\begin{aligned} (\tau_1, s_1) \star (\tau_2, s_2) &= [(\tau_1, 0) + (0, s_1)] \star [(\tau_2, 0) + (0, s_2)] \\ &= (\tau_1 \tau_2 + s_1 \cdot s_2, \tau_2 s_1 + \tau_1 s_2). \end{aligned} \quad (1)$$

In this way, the point set subject to a spacetime frame is an algebra (a slightly different perspective on this is given in the Discussion section). Note that this is distinct from other algebras that have been associated with relativistic spacetime [11-13], which are Clifford algebras and do not address the question of deriving the Minkowski norm from comparable first principles (this is another issue considered a bit further in the Discussion section). The magma-derived algebra we have identified above is commutative and nonassociative (as opposed to Clifford Algebras, which are associative and usually noncommutative), and is a Spin Factor Jordan algebra.

The *sum* operation of a vector space allows the definition of linear transformations, and the Euclidean norm implies linear isometries as given by the orthogonal group. In this case, equality of magnitude of two points of  $\mathbb{R}^4$  is invariant under the application of any member of  $O(n)$ . The symmetry group  $O(4)$  is not compelling for  $\mathbb{R} \oplus \mathbb{R}^3$ , since time and space axes do not necessarily seem interchangeable based on their respective envisioned dynamical roles (and the asymmetry suggested by the co-dimension 1 foliation of spacetime). However, the additional *product* operation of the spacetime magma-derived algebra allows recognition of an alternative constraint - the invariance of equality of magnitude of two points under (multiplicative) inversion rather than (orthogonal group) rotation-reflection. This constraint becomes quite attractive in light of inherent magma-derived algebra inversion symmetry, as noted below.

Explicitly, we make the following assumptions concerning the magnitude function  $f$  on the entirety of the spacetime magma-derived algebra:

1. *Positivity on the units.*  $f : \mathbb{R} \oplus \mathbb{R}^3 \rightarrow \mathbb{R}$  is such that  $f(x) > 0$  if  $x$  is a unit of the algebra (i.e., if  $x^{-1}$  exists),
2. *Homogeneity.* For  $\alpha \in \mathbb{R}$ ,  $f(\alpha x) = |\alpha| f(x)$ ,
3.  *$O(3)$ -invariance.*  $f$  is invariant under an orthogonal transformation of  $\mathbb{R}^3$ , i.e.,  $f(\tau, s) = f(\tau, g[s])$ , for  $g \in O(3)$  with  $g[s]$  being the point of  $\mathbb{R}^3$  that  $s$  is mapped to under the action of  $g$ ,
4. *Magnitude of the identity element.*  $f(\mathbf{1}) = \|\mathbf{1}\|$ , where  $\mathbf{1}$  is the identity element of the magma-derived algebra (so that  $\mathbf{1} = (1, 0)$  in this case), and  $\|\cdot\|$  is the Euclidean norm (so that  $\|\mathbf{1}\| = 1$ ),
5. *Equality of magnitude of units implies equality of magnitude of their inverses.* For any units  $x_1, x_2$  in the magma-derived algebra, if  $f(x_1) = f(x_2)$  then  $f(x_1^{-1}) = f(x_2^{-1})$ .

The above assumptions imply that  $f$  is the Minkowski norm. That is, Positivity and Homogeneity imply that for any unit  $x$  in the algebra there exists  $\alpha > 0$  such that  $f(x) = \alpha f(\mathbf{1}) = f(\alpha \mathbf{1})$ . "Invariance of equality of magnitude under inversion" (assumption 5 above) then implies that for any unit of the magma-derived algebra,

$$f(x)f(x^{-1}) = f(\alpha \mathbf{1})f(\alpha^{-1} \mathbf{1}) = [f(\mathbf{1})]^2 = \|\mathbf{1}\|^2. \quad (2)$$

Thus,  $f(x^{-1}) = \|\mathbf{1}\|^2 / f(x)$ , and so inversion of a point is accompanied by inversion of its magnitude. Since the equation  $f(x^{-1})f(x) = \|\mathbf{1}\|^2$  is invariant under the transformation  $x \rightarrow x^{-1}$ , we consider that equation to express "inversion symmetry". On our magma-derived algebra the inverse of a unit is evidently given by

$$(\tau, s)^{-1} = \frac{(\tau, -s)}{\tau^2 - s \cdot s}. \quad (3)$$

Using (2), the Homogeneity property, and the Euclidean  $\mathbb{R}^3$  reflection symmetry (i.e.,  $f(\tau, s) = f(\tau, -s)$ , from assumption 3 above), we then have

$$f(\tau, s)f((\tau, s)^{-1}) = f(\tau, s)f\left(\frac{(\tau, -s)}{\tau^2 - s \cdot s}\right) = \frac{[f(\tau, s)]^2}{|\tau^2 - s \cdot s|} = \|\mathbf{1}\|^2 = 1,$$

and so  $f(x)$  is the Minkowski norm (and the latter might be more properly labeled a magmitude rather than a magnitude since there are nonzero elements for which the Minkowski norm is zero).

The square of the Minkowski norm is a quadratic form on domains where  $\tau^2 - s \cdot s$  has a consistent sign, and the Minkowski inner product follows in the usual way that inner products are derived from quadratic forms. That is, we have the Polarization Identity: if  $Q(\cdot)$  is a quadratic form on a linear space, and  $w, z$  members of this space, then we define  $\langle w, z \rangle \equiv (Q(w+z) - Q(w) - Q(z))/2$ , where  $\langle w, z \rangle$  is then seen to be a symmetric bilinear form. Taking  $Q(x) = f^2(x)$ , we obtain the (pseudo-)inner product  $\langle (\tau_1, s_1), (\tau_2, s_2) \rangle = \tau_1 \tau_2 - s_1 \cdot s_2$ , which defines the Minkowski metric. Alternatively, the Minkowski inner product arises spontaneously from the quotient of two points, as is clear from (4) below (and more perspective on this is given in the first subsection of the Discussion section) - so it is not strictly necessary to apply the Polarization Identity.

Thus, one may conclude that Minkowski spacetime is a consequence of an inversion symmetry stemming from a magma implied by the inherent products linking  $\mathbb{R}$  and Euclidean  $\mathbb{R}^3$ , ultimately deriving from identification of an observer's time and space with the latter vector spaces.

**Mathematical implications of the magma-derived algebra.** As alluded to just prior to the above enumeration of the assumptions of the derivation of the Minkowski norm, there are inherent features of our spacetime magma-derived algebra that relate to inversion symmetry, and serve to further justify assumption 5 of the above derivation of the Minkowski norm:

First,  $x^{-1}$  is an irrotational vector field. Equivalently, considered as a 1-form (covector field),  $x^{-1}$  is curl-free (or more precisely, the 1-form  $\langle x^{-1}, \cdot \rangle$  is curl-free, where  $\langle \cdot, \cdot \rangle$  is now the Euclidean inner product), and thus is the exterior derivative of a 0-form, the latter scalar function being the integral of  $x^{-1}$ , which on a simply connected domain comprising only units and containing

$\mathbf{1}$  is the path independent  $\int_1^x u^{-1} \cdot du$  up to a constant (where, unless specified otherwise, “ $\cdot$ ” is the Euclidean inner product, in this case four-dimensional).

Second, we have the equation  $x^{-1} \cdot x = \mathbf{1}$ , so a unit and its multiplicative inverse have a kind of reciprocity with respect to the Euclidean inner product. The appearance of this equation is unchanged by the transformation  $x \rightarrow x^{-1}$ , so we consider it to express “inversion symmetry”. Since that equation indicates that the inner product of a point with its inverse is constant over spacetime, two points cannot be distinguished based on their (Euclidean, i.e., with fully symmetric terms) inner products with their respective inverses. Since a magnitude function is a one-dimensional representation of features of the space (and should not add “extraneous information”), it would be expected that a magnitude function faithful to the algebra should also embody the above lack of distinction between units. That is, it should have the feature that points are not distinguishable from the product of their magnitude with the magnitude of their inverse,  $f(x^{-1})f(x) = 1$  (equation (2)), from which the Minkowski norm follows.

Now, because it is commutative, the above spacetime magma-derived algebra is trivially a  $*$ -algebra where the involution  $*$  is the identity (in the Discussion section, an alternative involution on this algebra is considered). This is important to recognize so that we can place the Minkowski norm in the context of analogous structures that occur more generally. So, to begin, we recall that a  $*$ -algebra is simply an algebra that is associated with an (anti-)involution. That is, in addition to the algebra  $A$ , there is an operation  $*$  :  $A \rightarrow A$  such that for  $a \in A$  we have  $(a^*)^* = a$ , and for  $a, b \in A$  we have  $(ab)^* = b^*a^*$ . Familiar examples of  $*$ -algebras include any commutative algebra (trivially, with  $*$  being the identity), the complex numbers (where the involution could be either the identity or the complex conjugate operation), the quaternions (where  $*$  is a generalization of the complex conjugate), and the algebras of real or complex matrices (with  $*$  being the transpose or Hermitian transpose, respectively). Clifford algebras also become  $*$ -algebras once a particular (anti-)involution is selected.

In the context of the magma-derived algebra’s role as a  $*$ -algebra, the two inversion features noted in the two paragraphs before last then translate into assertions that  $x^{-*} \cdot x = \|\mathbf{1}\|^2$ , and  $\int_1^x u^{-*} \cdot du$  is path independent on a simply connected domain of units containing  $\mathbf{1}$  and is thus the integral of curl-free  $x^{-*}$ . We will show that if the latter features hold for a particular real  $n$ -dimensional unital  $*$ -algebra then we can always obtain a positive-homogeneous magnitude function ( $f(\alpha x) = \alpha f(x)$  for  $\alpha > 0$ ) with level sets given by the level sets of the integral of (co)vector field  $x^{-*}$ . We refer to these functions as “inverse norms” in contrast to the Euclidean norm, a positive-homogeneous magnitude function with level sets given by the level sets of the integral of (co)vector field  $x$ . While the Euclidean norm solves the equation  $x \cdot x = f(x)f(x)$  with  $O(n)$ -symmetric left-hand-side, we shall see that an inverse norm solves the  $*$ -inverse-symmetric equation  $x^{-*} \cdot x = f(x^{-*})f(x) = \|\mathbf{1}\|^2$ . Examples of inverse norms include the Euclidean norm on the Cayley-Dickson algebras (the reals, the complex numbers, the quaternions, the octonions, the sedenions, ...), the Minkowski norm on the Spin Factor Jordan algebras, and the  $m$ -th root of the determinant multiplied by  $\sqrt{m}$  on the algebra of real  $m \times m$  matrices. From this standpoint, the latter three structures are the same thing expressed in the context of different algebras.

Inverse norms share with Euclidean norms the feature that almost all points can be decomposed as the product of a magnitude function with a gradient of the magnitude function, where the gradient of the magnitude function itself has unit magnitude (if  $f(x)$  is taken to be the Euclidean norm, then  $x \neq 0$  is decomposed as  $x = f(x)\nabla f(x)$  with  $f(\nabla f(x)) = 1$ ; for inverse norms, see Theorem 3 below).

Before placing the Minkowski norm in the more general context of inverse norms, we will consider the physical meaning of the magma-derived spacetime algebra.

**Physical implications of the magma-derived algebra.** As was our intent, we have derived the Minkowski norm independent of usual physical arguments - i.e., we have used only the entities of time and space, which can also be thought of in purely mathematical terms (as justified by physical experiments). Once the Minkowski norm is derived, it has (of course) well known profound physical implications such as the relativity of simultaneity, the contraction of length in the direction of velocity, and (once energy and momentum are introduced) the equivalence of mass and energy. However, although we have forwarded a derivation that is ‘prior to’ subsequently introduced physics (of the type cited in the Introduction), the mathematical structure underlying this derivation has intrinsic physical meaning.

Velocity is one of the most basic of physical concepts, and it is expressed as a quotient of quantities related strictly to space (in the numerator) and strictly to time (in the denominator). From a general perspective, this structure in isolation could be considered a little unsatisfactory (or incomplete): If from an observer’s viewpoint we divide a space quantity  $s \in \mathbb{R}^3$  by a time quantity  $\tau \in \mathbb{R}$  to obtain a velocity  $s/\tau$ , we know from aforementioned isomorphisms that this can be equivalently thought of as point  $(0, s)$  on the frame’s space component divided by point  $(\tau, 0)$  on the frame’s time component. The magma product (pre-existing on the union of the frame components) then provides the agreeable  $(0, s) \star (\tau, 0)^{-1} = (0, s/\tau)$  as an analogue of the previously computed velocity  $s/\tau$ . But how is it that only points residing on frame components are able to be divided? A general perspective recognizing the validity of all inertial frames “demands” that we be able to divide any spacetime unit by any other spacetime unit within a specific spacetime frame (otherwise it must be conceded that frame component points are very special, while a default viewpoint would suggest they are not so special). Indeed, the physical concept of velocity as a quotient of points on respective frame components motivates us to not only supply the implied division process for general spacetime points (that are units), but also to do this in a manner that is Lorentz covariant. In fact, this is precisely what the inherent spacetime magma-derived algebra accomplishes. Recalling (1) and (3), if we divide  $(\tau_1, s_1)$  by  $(\tau_2, s_2)$ , we obtain

$$(\tau_1, s_1) \star (\tau_2, s_2)^{-1} = \frac{(\tau_1 \tau_2 - s_1 \cdot s_2, \tau_2 s_1 - \tau_1 s_2)}{\tau_2^2 - s_2 \cdot s_2}. \quad (4)$$

On the right-hand-side above, the denominator and the two components of the numerator are evidently Lorentz covariant expressions. Thus, the act of computation of the quotient of two points is Lorentz covariant (however, the *result* of the computation is a single point which, of course, is not Lorentz covariant - e.g., velocity is relative to a frame).

Furthermore, the operation of division in the magma-derived spacetime algebra has a definite physical meaning apart from its Lorentz covariant treatment of the ratios from which a velocity is

computed. The first component of the ordered pair on the right-hand-side of (4) is the Minkowski inner product of the two points - reflecting the deviation of the two points from being pseudo-orthogonal (points that are pseudo-orthogonal lie on the respective components of a possible spacetime frame). In contrast, the second component of the ordered pair reflects the deviation of the two points from being linearly dependent. If the two points are linearly dependent, then there is an inertial frame such that they and the origin of the spacetime frame are events that occur at the same point in space in the context of that inertial frame. Note that this second component is antisymmetric and bilinear. Analogous to outer products in Clifford algebras, this latter expression can be taken to be the *Minkowski outer product*. This perspective is further developed in the Discussion section. At this point, we summarize these physical aspects as follows:

- The magma-derived algebra satisfies the need to provide a Lorentz covariant expression for the operation of division of spacetime units - an implied generalization of the procedure of computing a velocity.
- The output of the magma-derived algebra's division operation is a point with complementary relativistic components - the first being proportional to the Minkowski inner product, and the second being proportional to the Minkowski outer product - these having manifestly evident physical interpretations.

Finally, we also note that the Lorentz covariant expression for velocity given by the right-hand-side of (4) itself provides an alternative route for derivation of the Minkowski metric from spacetime magma, and that route will be explored elsewhere.

### The Minkowski norm in context: Inversion symmetry and inverse norms on unital \*-algebras

It is now our objective to formalize the understanding of the Minkowski norm as an example of a structure that occurs as a fundamental feature in a variety of mathematical contexts that are both familiar and have seemingly little in common with Special Relativity. The implication will be that it is not at all surprising that the Minkowski norm arises in spacetime.

We have previously noted that

- The Euclidean norm on  $\mathbb{R}^n$  is a positive-homogeneous magnitude function  $f(x)$  with level sets given by the level sets of the integral of (co)vector field  $x$ , and which solves the equation  $x \cdot x = f(x)f(x)$  with  $O(n)$ -symmetric left-hand-side.
- In the context of our magma-derived spacetime \*-algebra, the Minkowski norm is a positive-homogeneous magnitude function  $f(x)$  with level sets given by the level sets of the integral of (co)vector field  $x^{-*}$ , and which solves the \*-inverse-symmetric equation  $x^{-*} \cdot x = f(x^{-*})f(x) = \|1\|^2$ .
- Two features of the magma-derived algebra are that on an appropriate domain of units we have  $x^{-*} \cdot x = \|1\|^2$ , and  $x^{-*}$  is irrotational. The latter express the “inversion symmetry” of the algebra, and are precisely the features needed for the existence of the important analogues of the Minkowski norm on other \*-algebras.

Our use of the term “inversion symmetry” refers to a discrete symmetry. That is, the relevant group action leaves an equation  $x^{-*} \cdot x = f(x^{-*})f(x) = \|1\|^2$  unchanged for each  $x$  in a domain of units. The group in question has two elements: the identity ( $x \rightarrow x$ , for each  $x$  in the domain), and the operation of \*-inversion ( $x \rightarrow x^{-*}$ , for each  $x$  in the domain). In contrast, it is the *solution* to

the equation over the specified domain of units (the inverse norm) that is expected to be invariant under the action of a Lie group.

To construct a magnitude function on a unital \*-algebra mimicking the Euclidean norm but based on vector field  $x^{-*}$  rather than vector field  $x$ , it is first required that  $x^{-*}$  be a gradient, since the level sets of the positive-homogeneous magnitude function are to be determined by the level sets of the integral of that (co)vector field. To embody the property of “invariance of equality of magnitude under \*-inversion”, it is necessary for \*-inversion to map magnitude function level sets to magnitude function level sets. Equivalently, given that the magnitude function is positive-homogeneous, the requirement is that the magnitude function level set containing **1** be invariant under \*-inversion of the space. The latter also imply that the magnitude function values of an element and its \*-inverse are reciprocal.

Formalizing these statements, we have

**Definition 1.** A real finite dimensional unital \*-algebra has inversion symmetry if on some neighborhood of **1** the vector field  $x^{-*}$  is a gradient and

$$x^{-*} \cdot x = \|1\|^2. \quad (5)$$

The inversion symmetry is further characterized by

**Theorem 1.** A real unital \*-algebra has inversion symmetry if and only if there is a simply connected neighborhood  $D$  of **1** where a function  $f : D \rightarrow \mathbb{R}$  with  $\nabla f(x) = x^{-*}$  is such that the set of level sets of  $\hat{f}(x)$  and the set of level sets of  $\hat{f}(x^{-*})$  are the same set.

**Definition 2.** For a function  $\hat{f}$  and neighborhood  $D$  of **1** relevant to an algebra with inversion symmetry as in Theorem 1, an inverse norm is a positive-homogeneous function  $f : D \rightarrow \mathbb{R}^+$  having the same level sets as  $\hat{f}$  and such that

$$f(x^{-*}) = \frac{\|1\|^2}{f(x)}. \quad (6)$$

**Theorem 2.** An algebra with inversion symmetry has a unique inverse norm.

The above theorems are corollaries of the following, which provides another characterization of an inverse norm.

**Theorem 3.** A real finite dimensional unital \*-algebra has inversion symmetry if and only if on this algebra there exists a simply connected domain  $D$  containing **1** with  $x \in D$  implying both  $x^{-*} \in D$  and  $\alpha x \in D$  for any real  $\alpha > 0$ , such that there exists  $f : D \rightarrow \mathbb{R}^+$ , with

- $f(1) = \|1\|$ ,
- $f(x_1) = f(x_2)$  implies  $f(x_1^{-*}) = f(x_2^{-*})$ ,
- any element of  $D$  can be decomposed as

$$x = f(x)\nabla f(x^{-*}), \quad (7)$$

with

$$f(\nabla f(x^{-*})) = 1, \quad (8)$$

where  $\nabla f(x^{-*})$  indicates that the gradient  $\nabla f$  is evaluated at the point  $x^{-*}$ .

**Corollary 1.** The function  $f$  specified by Theorem 3 is the inverse norm associated with the given algebra.

In summary: a \*-algebra on which  $x^{-*}$  is an irrotational vector field, has inversion symmetry if the expression  $x^{-*} \cdot x$  is invariant,

i.e., constant. This constant must necessarily be  $\|\mathbf{1}\|^2$  considering the case of  $x = \mathbf{1}$ . Since the expression  $x^{-*} \cdot x = \|\mathbf{1}\|^2$  doesn't change under the transformation  $x \rightarrow x^{-*}$ , we refer to that property as expressing inversion symmetry. An inverse norm  $f$  on the algebra is required to be positive homogenous (like norms in general), but be developed from the integral of  $x^{-*}$  (in contrast to the Euclidean norm, which is positive homogeneous but develops from the integral of  $x$ ), and it must solve the  $*$ -inverse symmetric functional equation,  $f(x^{-*})f(x) = \|\mathbf{1}\|^2$ . For the magma-derived algebra of the last section,  $f$  is then the Minkowski norm. But the inverse norm turns out to be an important structure in the setting of other algebras having nothing to do with Special Relativity. Understanding of Special Relativity is deepened by recognizing its underlying kinship with these different areas, and that is the reason we have introduced the above definitions and theorems. For example, when considered from the standpoint of its inherent algebraic structure as ultimately expressed by the inverse norm, the magma-derived algebra of Special Relativity is kin to any of the Cayley-Dixon algebras, and indeed to the  $m \times m$  real matrix algebra itself. From that perspective, the appearance of the Minkowski norm in spacetime is not strange at all, given that the physical entity of velocity suggests (as we have argued) that spacetime should be viewed as an algebra.

Thus, inversion symmetry (Definition 1) is the feature of the magma-derived algebra which implies that a spacetime frame has a Minkowski norm - the prototype of an inverse norm. More generally, inversion symmetry is the thing that is necessary and sufficient for an inverse norm to be defined on an algebra, and is thereby responsible for recognition of a revealing structural similarity between Special Relativity and other seemingly remote mathematical settings.

## Proof of the Theorems

As an aid in identifying the key steps in the proof of Theorem 3, we will italicize certain sentences. A general idea of the strategy of the proof can be obtained by initially concentrating on these italicized sentences.

**Proof of Theorem 3, Necessity.** For the Necessity portion of the proof, inversion symmetry is assumed, and we must show that it implies all statement following the "if and only if" clause in the statement of the theorem.

*Inversion symmetry implies existence of a simply connected domain containing  $\mathbf{1}$  on which  $x^{-*}$  is a gradient and  $x^{-*} \cdot x = \|\mathbf{1}\|^2$ . Without loss of the prior two properties, this domain can be enlarged to meet the specifications of  $D$  in the theorem statement.* That is, we first enlarge the original domain to consist of all non-origin points lying on the rays from the origin through any point of the original domain. This new neighborhood  $\hat{D}$  is still simply connected and it is clear that  $x^{-*}$  is a gradient and  $x^{-*} \cdot x = \|\mathbf{1}\|^2$  on  $\hat{D}$ . Consider  $\hat{D}^{-*} \equiv \{x^{-*} : x \in \hat{D}\}$ . The operation  $x \rightarrow x^{-*}$  maps open rays to open rays. Since that operation is a diffeomorphism between  $\hat{D}$  and  $\hat{D}^{-*}$ , the set  $\hat{D}^{-*}$  is simply connected, is a union of open rays, and shares at least one of these rays with  $\hat{D}$  (the ray through  $\mathbf{1}$ ). It follows that  $D \equiv \hat{D} \cup \hat{D}^{-*}$  is simply connected, and  $x \in D$  implies both  $x^{-*} \in D$  and  $\alpha x \in D$  for  $\alpha > 0$ .

*Thus, there exists a function  $\hat{f}$  on  $D$  such that*

$$x^{-*} = \nabla \hat{f}(x). \quad (9)$$

*Selection of a particular constant of integration fixes  $\hat{f}$ . Since  $(\alpha x)^{-*} = \alpha^{-1} x^{-*}$ , applying (9) twice yields*

$$(\alpha x)^{-*} = \nabla \hat{f}(\alpha x) = \alpha^{-1} \nabla \hat{f}(x) = \alpha^{-1} x^{-*}. \quad (10)$$

Now consider  $D = D_1 \cup D_2$  where  $D_1$  is the subset of  $D$  whose members can be written as a positive multiple of some member of the level set of  $\hat{f}$  that contains  $\mathbf{1}$ , and  $D_2$  is its complement in  $D$ . If  $D_2$  is nonempty, then since  $D$  is connected there is an  $w \in D$  such that every neighborhood of  $w$  contains a point of  $D_1$  and a point of  $D_2$ . The ray from the origin through  $w$  must necessarily be tangent to the level set of  $\hat{f}$  containing  $\mathbf{1}$  at some point  $y \in D$  (otherwise, there either won't be rays from the origin passing arbitrarily near  $x$  that intersect the level set containing  $\mathbf{1}$  or there won't be rays arbitrarily near  $w$  not intersecting the level set containing  $\mathbf{1}$ ). But then the gradient of  $\hat{f}$  at  $y$  is normal to the ray from the origin passing through  $y$ . Since  $\nabla \hat{f}(y) = y^{-*}$ , this means that  $y^{-*} \cdot y = 0$ , which contradicts (5). Since  $D_1$  cannot be empty (it contains the open ray from the origin containing  $\mathbf{1}$ ), it follows that  $D_2$  is empty. Thus, any  $x \in D$  can be written as  $x = \alpha z$  for some  $\alpha > 0$  and some  $z$  in the level set of  $\hat{f}$  containing  $\mathbf{1}$ . We can also see that this decomposition is unique. That is, suppose for some  $x \in D$  we have  $x = \alpha_1 z_1 = \alpha_2 z_2$  with  $z_1, z_2$  distinct points on the level set of  $\hat{f}$  containing  $\mathbf{1}$ . Since the gradient vectors  $\nabla \hat{f}$  are parallel to each other at each point of the open ray from the origin through  $s$  (as is evident from (10)),  $\hat{f}$  is strictly increasing, strictly decreasing, or constant, along the open ray. And yet,  $x = \alpha_1 z_1 = \alpha_2 z_2$  implies that the ray intersects the level set containing  $\mathbf{1}$  at least twice, which is a contradiction if  $\hat{f}$  is either strictly increasing or strictly decreasing. Thus,  $\hat{f}$  would be required to be constant along the open ray containing  $x$ , implying that the gradient at  $x$  (which is  $\nabla \hat{f}(x) = x^{-*}$ ) is normal to this ray - and so its inner product with  $x$  is zero, again contradicting (5). Thus, the decomposition  $x = \alpha z$  is unique.

For a fixed choice of  $\alpha > 0$ ,

$$\{\alpha z : z \text{ is a member of the level set of } \hat{f} \text{ containing } \mathbf{1}\} \quad (11)$$

is a level set of  $\hat{f}$ , because the second equality in (10) implies that the gradient of  $\hat{f}$  at any point  $\alpha z$  of the set (11) is proportional to the gradient at the corresponding open ray point  $z$  on the level set containing  $\mathbf{1}$ . Thus, the level sets of  $\hat{f}$  are each a dilation of the level set containing  $\mathbf{1}$ .

We now construct a positive-homogeneous function  $f : D \rightarrow \mathbb{R}$  which will have the same level sets as  $\hat{f}$ . First, we specify that all points  $z$  in the level set of  $\hat{f}$  containing  $\mathbf{1}$  are given the value  $f(z) = \|\mathbf{1}\|$ . As we have seen, any point  $s \in D$  can be uniquely written as  $s = \alpha z$  with  $\alpha > 0$  and  $z$  in the level set of  $\hat{f}$  containing  $\mathbf{1}$ . We thus complete the definition of  $f$  on points of  $D$  via

$$f(x) = f(\alpha z) \equiv \alpha f(z) = \alpha \|\mathbf{1}\|. \quad (12)$$

The functions  $f$  and  $\hat{f}$  are seen to have the same level sets in  $D$ , since their level sets are dilations of the level set of  $\hat{f}$  containing  $\mathbf{1}$ .

Since by assumption  $x^{-*}$  is a gradient on  $D$ , the level set of  $\hat{f}$  (and  $f$ ) containing  $\mathbf{1}$  is smooth in  $D$ , and by construction  $f$  is differentiable on  $D$ . For  $x \in D$ , we claim that for  $\alpha > 0$ ,

$$\nabla f(\alpha x) = \nabla f(x). \quad (13)$$

This follows because the level sets of  $f$  are dilations of each other via (12), and so the set of directional derivatives of  $f$  at a point  $x$  will be proportional to the set of directional derivatives of  $f$  at a point  $\alpha x$ . But (12) implies that for any point on the open ray defined by the origin and any of the points  $z$  on the level set containing  $\mathbf{1}$ , the directional derivative of  $f$  with respect to a unit vector along the open ray is constant everywhere on the open ray (a nonzero constant, because of (5) and (9)) - thus, the other directional derivatives, and the gradient itself, must also be constant along the points of the ray, and (13) follows.

Now consider  $\nabla \hat{f}$  at points of  $D$  on the open ray defined by  $\mathbf{0}$  and  $\mathbf{1}$ . From (9), we have that the gradient of  $\hat{f}$  at points of this ray is directed along the ray. Since  $\hat{f}$  and  $f$  have the same level sets, it follows that  $\nabla \hat{f}$  at points on this same ray also is directed along this ray. The norm of  $\nabla \hat{f}$  at points on this ray is thus the value of the directional derivative specified by the vector of norm 1 in the direction of the ray. The unit norm vector in this direction is  $\mathbf{1}/\|\mathbf{1}\|$ . But  $f(\alpha \mathbf{1}/\|\mathbf{1}\|) = \alpha f(\mathbf{1})/\|\mathbf{1}\| = \alpha$ , which follows from (12) and our earlier specification of the value of points on the level set containing  $\mathbf{1}$ , where in particular,  $f(\mathbf{1}) = \|\mathbf{1}\|$ . Hence, the derivative of  $f$  in the direction given by unit vector  $\mathbf{1}/\|\mathbf{1}\|$  is 1, and we then have  $\nabla f(\mathbf{1}) = \mathbf{1}/\|\mathbf{1}\|$ . Equation (9) implies that  $\nabla \hat{f}(\mathbf{1}) = \mathbf{1}$ . Thus,  $\nabla \hat{f}(\mathbf{1}) = \|\mathbf{1}\| \nabla f(\mathbf{1})$ . Since the level sets of  $f$  are also level sets of  $\hat{f}$ , the gradients of  $f$  and  $\hat{f}$  must be proportional at each point of a level set. Hence, for  $z$  on the level set containing  $\mathbf{1}$ , the last equation implies

$$\nabla \hat{f}(z) = \|\mathbf{1}\| \nabla f(z). \quad (14)$$

So, for  $z$  on the level set containing  $\mathbf{1}$ , (12), (13), (14), and (9), imply that

$$\begin{aligned} f((\alpha z)^{-*}) \nabla f(\alpha z) &= \alpha^{-1} f(z^{-*}) \nabla f(z) \\ &= \frac{\alpha^{-1} f(z^{-*}) \nabla \hat{f}(z)}{\|\mathbf{1}\|} \\ &= \frac{f(z^{-*})}{\|\mathbf{1}\|} (\alpha z)^{-*}. \end{aligned} \quad (15)$$

We will now show that  $f(z^{-*}) = \|\mathbf{1}\|$  for any  $z$  in the level set containing  $\mathbf{1}$ , from which (15) will imply (7). First, from (9), for  $u \in D$ ,

$$\hat{f}(u) - \hat{f}(\mathbf{1}) = \int_{\mathbf{1}}^u \nabla \hat{f}(x) \cdot dx = \int_{\mathbf{1}}^u x^{-*} \cdot dx, \quad (16)$$

since  $D$  is simply connected. Similarly, since by assumption  $x \in D$  implies  $x^{-*} \in D$ ,

$$\hat{f}(u^{-*}) - \hat{f}(\mathbf{1}) = \int_{\mathbf{1}}^{u^{-*}} \nabla \hat{f}(x) \cdot dx = \int_{\mathbf{1}}^{u^{-*}} x^{-*} \cdot dx \quad (17)$$

$$= \int_{\mathbf{1}}^u x \cdot d(x^{-*}) \quad (18)$$

$$= u \cdot u^{-*} - \|\mathbf{1}\|^2 - \int_{\mathbf{1}}^u x^{-*} \cdot dx \quad (19)$$

$$= -(\hat{f}(u) - \hat{f}(\mathbf{1})), \quad (20)$$

where (18) follows from the change of variable  $x \rightarrow x^{-*}$ , (19) follows from integration by parts, and (20) follows from (16) and (5).

Equation (20) implies that  $u \in D$  is on the level set of  $\hat{f}$  containing  $\mathbf{1}$  if and only if  $u^{-*}$  is on the level set of  $\hat{f}$  containing  $\mathbf{1}$ . Since in  $D$  the level set of  $f$  containing  $\mathbf{1}$  is the same as the level set of  $\hat{f}$  containing  $\mathbf{1}$ , we then have that  $u$  is on the level set of  $f$  containing  $\mathbf{1}$  if and only if  $u^{-*}$  is on the level set of  $f$  containing  $\mathbf{1}$ . This means that the term multiplying  $(\alpha z)^{-*}$  on the right-hand-side of (15) is unity (since we now have  $f(z^{-*}) = \|\mathbf{1}\|$  because  $z$  is on the level set containing  $\mathbf{1}$  which is now seen to imply that  $z^{-*}$  is on the level set containing  $\mathbf{1}$  and we have previously defined  $f$  to have the value  $\|\mathbf{1}\|$  on the level set containing  $\mathbf{1}$ ). As we have seen earlier, any  $s \in D$  can be written uniquely as  $s = \alpha z$  for some  $z$  on the level set of  $f$  containing  $\mathbf{1}$  and some  $\alpha > 0$ . Consequently, (15) implies that for  $x \in D$ ,

$$x^{-*} = f(x^{-*}) \nabla f(x), \quad (21)$$

which is equivalent to (7). Applying  $f$  to both sides of (7) and then using the second equality in (12) we obtain (8).

Finally, (7) and (21) imply

$$x \cdot d(x^{-*}) + x^{-*} \cdot dx = f(x) \nabla f(x^{-*}) \cdot d(x^{-*}) + f(x^{-*}) \nabla f(x) \cdot dx, \quad (22)$$

which is equivalent to

$$d(x^{-*} \cdot x) = d(f(x^{-*}) f(x)). \quad (23)$$

Since by assumption  $x^{-*} \cdot x = \|\mathbf{1}\|^2$ , the left-hand-side above is zero, and so  $f(x^{-*}) f(x)$  is constant on  $D$ . This means that for  $x_1, x_2 \in D$  we have  $f(x_1^{-*}) f(x_1) = f(x_2^{-*}) f(x_2)$ . Thus  $f(x_1) = f(x_2)$  implies  $f(x_1^{-*}) = f(x_2^{-*})$ . Q.E.D.

**Proof of Theorem 3, Sufficiency.** For the Sufficiency portion of the proof, all of the statements following the theorem's "if and only if" clause are assumed, and inversion symmetry must be derived from them - i.e., it must be shown that  $x^{-*}$  is a gradient on  $D$  and  $x^{-*} \cdot x = \|\mathbf{1}\|^2$ .

Equation (7) implies (21) and together these imply (22) and (23). Integrating both sides of (23) on any path in simply connected  $D$  from  $\mathbf{1}$  to  $u \in D$ , and noting that by assumption  $\|\mathbf{1}\|^2 = (f(\mathbf{1}))^2$ , we obtain

$$u^{-*} \cdot u = f(u^{-*}) f(u). \quad (24)$$

Since  $f$  is strictly positive on  $D$  and  $x \in D$  implies  $x^{-*} \in D$ , this means that  $x^{-*} \cdot x \neq 0$ .

Now we establish positive-homogeneity, i.e., for  $\alpha > 0$ ,  $f(\alpha x) = \alpha f(x)$ . Applying (7) twice, we have

$$\alpha x = f(\alpha x) \nabla f(\alpha^{-1} x^{-*}) = \alpha f(x) \nabla f(x^{-*}). \quad (25)$$

Equation (8) then gives

$$f((\alpha x)/f(\alpha x)) = f(x/f(x)) = 1. \quad (26)$$

If  $f(\alpha x) \neq \alpha f(x)$  then (26) implies that there are two different points on the ray from zero through  $x$  having the same magnitude. But according to (25),  $\nabla f(x^{-*})$  and  $\nabla f(\alpha^{-1}x^{-*})$  are parallel, and so  $\nabla f(x)$  and  $\nabla f(\alpha x)$  are parallel for any  $x \in D$ . This means that for points of  $D$  on the ray through  $x$  it must be that  $f$  is strictly increasing or strictly decreasing or constant. Since there are two points on the ray having the same value of  $f$ , it follows that  $f$  must be constant on the ray, so that  $\nabla f(x)$  is normal to the ray from  $0$  through  $x$ , i.e.,  $x \cdot \nabla f(x) = 0$ . But from (21),  $x \cdot x^{-*} = f(x^{-*})(s \cdot \nabla f(x))$  and we have already noted that (24) indicates that the left-hand-side of this equation is nonzero, contradicting the prior sentence. Thus, it is required that  $f(\alpha x) = \alpha f(x)$ .

Since by assumption  $f(x) > 0$  in  $D$ , for any particular choice of  $x \in D$  we can find an  $\alpha > 0$  such that  $f(x) = \alpha \|1\| = \alpha f(1) = f(\alpha 1)$ , where the last equality follows from positive-homogeneity. Since the value of  $f$  is thus the same at  $s$  and  $\alpha 1$ , by assumption we have  $f(s^{-*}) = f(\alpha^{-1}1)$ . Once again using positive-homogeneity, we obtain the functional equation

$$f(x^{-*})f(x) = f(\alpha^{-1}1)f(\alpha 1) = (f(1))^2 = \|1\|^2. \quad (27)$$

Equation (5) then follows from (24).

We now only need show that  $x^{-*}$  is a gradient on  $D$ . Equations (7) and (8) indicate that every  $x \in D$  can be written as  $x = \beta w$  with  $\beta > 0$  and  $w$  in the level set of  $f$  with the value 1. This decomposition must be unique because otherwise we would have  $x = \beta_1 w_1 = \beta_2 w_2$  with  $w_1, w_2$  on the level set of  $f$  with value 1. But after applying  $f$  to both sides of the second equality in the last sentence, positive-homogeneity would imply  $\beta_1 = \beta_2$ , which would then imply  $w_1 = w_2$ . By positive-homogeneity, the level set of  $f$  containing  $1$  is just a dilation of the level set having value 1. Thus, we can uniquely write  $x = \alpha z$  with  $\alpha > 0$  and  $z$  on the level set of  $f$  containing  $1$  (the points of this set have value  $\|1\|$  by assumption). Hence, we can define a new function  $\hat{f}(x) = \hat{f}(\alpha z) \equiv \|1\|(\log \alpha)f(z) = \|1\|^2 \log \alpha$ . Note that  $\hat{f}$  has the same level sets as  $f$ , so the gradients of  $\hat{f}$  and  $f$  are proportional at the points of any level set. Directional derivatives along the ray  $\{\alpha z : \alpha > 0\}$  result from  $df(\alpha z)/d\alpha = f(z) = \|1\|$  and  $d\hat{f}(\alpha z)/d\alpha = \|1\|^2/\alpha$  (using positive-homogeneity for the first equation and our above specification of  $\hat{f}$  for the last equation). Thus,  $\nabla \hat{f}(\alpha z) = (\|1\|/\alpha)\nabla f(\alpha z)$ . Since (27) implies that  $f(z^{-*}) = \|1\|$ , we have that (21) (equivalent to (7)), positive-homogeneity, and the equation of the last sentence, imply

$$x^{-*} = (\alpha z)^{-*} = f((\alpha z)^{-*})\nabla f(\alpha z) = \nabla \hat{f}(\alpha z) = \nabla \hat{f}(x).$$

The above indicates that  $x^{-*}$  is a gradient. *Q.E.D.*

**Proof of Theorem 1.** As regards *Sufficiency*, by assumption there is a function  $\hat{f}$  on a simply connected neighborhood  $D$  of  $1$  whose gradient is  $x^{-*}$ , i.e.,  $x^{-*}$  is a gradient on  $D$ . Since it is also assumed that  $\hat{f}(x)$  and  $\hat{f}(x^{-*})$  have the same set of level sets,  $x \in D$  implies  $x^{-*} \in D$ . Equations (16) through (19) then follow on this domain. Let  $Z$  be the set of points on the level set of  $\hat{f}$  containing  $1$ . For  $u = z \in Z$ , (16) implies that the integral on the right-hand-side

of (19) is zero. Since  $\hat{f}(x)$  and  $\hat{f}(x^{-*})$  have the same set of level sets, the level set of  $\hat{f}(x)$  containing  $1$  (the set  $Z$ ) is the same as the level set of  $\hat{f}(x^{-*})$  containing  $1$ . This implies that for  $u \in Z$  the left-hand-side of (17) is zero. Thus, for  $z \in Z$ , (19) implies  $z^{-*} \cdot z = \|1\|^2$ . Since  $\nabla \hat{f}(1)$  is normal to  $Z$  at  $1$ , there is a neighborhood  $D'$  of  $1$  such that every point of this neighborhood is on a ray from the origin through some point  $z \in Z$ , and so each  $s \in D'$  can be written as  $x = \alpha z$  with  $\alpha > 0$ ,  $z \in Z$ . We then have for any  $x \in D'$ ,  $x^{-*} \cdot x = z^{-*} \cdot z = \|1\|^2$ . Thus,  $D'$  is the neighborhood of  $1$  required by Definition 1 for inversion symmetry to be present.

As regards *Necessity*, inversion symmetry implies all of the statements of Theorem 3 following the “if and only if” clause. This implies equations (16) through (20), which imply that the set of level sets of  $\hat{f}(x)$  and the set of level sets of  $\hat{f}(x^{-1})$  are the same on  $D$ . *Q.E.D.*

**Proof of Theorem 2.** Inversion symmetry implies the statements of Theorem 3 following the “if and only if” clause (since the latter theorem has been proved), including the requisite simply connected domain  $D$  of that theorem. Therefore, we may use the arguments from the Proof of Sufficiency portion of the proof of Theorem 3, where it was shown that  $f$  is a positive-homogeneous function satisfying (27), i.e., (6). The final paragraph of that Proof of Sufficiency demonstrates the required relationship of  $f$  to a function  $\hat{f}$  with  $\nabla \hat{f}(x) = x^{-*}$ . Thus,  $f$  fulfills the requirements of Definition 2.

The level sets of the integral of  $x^{-*}$  are unique regardless of the constant of integration. These are also the level sets of the inverse norm, whose level set containing  $1$  is assigned the value  $\|1\|$ . The values of the remainder of the level sets of the inverse norm are then uniquely determined by the positive-homogeneity condition. Thus, the inverse norm is unique. *Q.E.D.*

The above proof of Theorem 2 also establishes Corollary 1.

## Discussion

### Modifications of the magma-derived spacetime $^*$ -algebra

While the product of two spacetime points (equation (1)) gives no immediate hint of what magnitude function should be associated with the magma-derived algebra, the inverse operation (division) does, as is clear from our remarks following the appearance of equation (4). Furthermore, the numerator on the right-hand-side of (4) by itself is a bilinear product on  $\mathbb{R} \oplus \mathbb{R}^3$ ,

$$(\tau_1, s_1) \hat{\star} (\tau_2, s_2) \equiv (\tau_1, s_1) \star (\tau_2, -s_2) = (\tau_1 \tau_2 - s_1 \cdot s_2, \tau_2 s_1 - \tau_1 s_2), \quad (28)$$

defining an algebra with a product given by the differences (rather than the sums as in (1)) of all magma-generating products across components of the two spacetime points. This new algebra is nonunital, noncommutative, nonassociative, and is not a  $^*$ -algebra. Essentially, it is a magma on the entirety of  $\mathbb{R} \oplus \mathbb{R}^3$  that is consistent with the latter’s linear structure. But this new magma has some nice features. The first component of the product on the right-hand-side of (28) is a symmetric bilinear form, and so we can define an inner product as

$$\langle (\tau_1, s_1), (\tau_2, s_2) \rangle \equiv \tau_1 \tau_2 - s_1 \cdot s_2,$$

which is, of course, the Minkowski inner product. The second component of the product on the right-hand-side of (28) is antisymmetric and bilinear, and so we can define an outer product as

$$\langle (\tau_1, s_1), (\tau_2, s_2) \rangle \equiv \tau_2 s_1 - \tau_1 s_2,$$

which we call the *Minkowski outer product*. Thus,

$$(\tau_1, s_1) \star (\tau_2, s_2) = \langle (\tau_1, s_1), (\tau_2, s_2) \rangle + \langle (\tau_1, s_1), (\tau_2, s_2) \rangle,$$

and

$$(\tau, s)^2 = \langle (\tau, s), (\tau, s) \rangle = \tau^2 - s \cdot s.$$

The latter implies a magnitude function (or magnitude function) as given by the Minkowski norm. Thus, this magma-algebra (i.e., a nonunital non- $\star$ -algebra with a nonassociative and noncommutative product) has a Lorentz invariant product, implies a “Minkowski outer product”, and implies a Minkowski metric. We have already noted that the first component of the product is the deviation of two points from being pseudo-orthogonal, and the second component expresses the deviation of two points from being linearly dependent - a juxtaposition of complementary geometric features of the two points considered in tandem.

We can also observe that the involution selected for the original magma-derived algebra (with product (1)) need not be the identity. Thus, we also have a  $\star$ -algebra if the involution is defined by  $(\tau, s)^* \equiv (\tau, -s)$ . It is easy to verify that the Minkowski norm still satisfies (5) through (8) given the understanding that the inner product and gradient are consistent with this choice of  $\star$ . That is, satisfaction of (5) now requires that “ $\cdot$ ” be the Minkowski inner product, i.e.,  $a \cdot b = a^\mu b_\mu$  where the contravariant vector is determined by a metric tensor implied by the involution. Similarly, (7) and (8) require that “ $\nabla$ ” be the four-gradient  $\partial^\mu$ . This suggests a liberalization of Definition 1 wherein the choice of inner product is made such that (5) holds. Nevertheless, we are also free to suggest that the involution be chosen such that the latter equation holds with the maximally symmetric Euclidean inner product. In either case, using the above alternative involution we are at least able to write

$$(\tau, s) \star (\tau, s)^* = (\tau^2 - s \cdot s) \mathbf{1},$$

while retaining the structure of a commutative unital  $\star$ -algebra.

The above alternative product and involution choices notwithstanding, it should be clear that the original magma-derived spacetime algebra can be understood as *inherent* in the vector space  $\mathbb{R} \oplus \mathbb{R}^3$  when  $\mathbb{R}^3$  is Euclidean and  $\mathbb{R}, \mathbb{R}^3$  are distinguished (as opposed to  $\mathbb{R}^2, \mathbb{R}^2$ ) by their respective identification with an observer’s “time” and “space” - the single underlying physical assumption. An analogy with vector space  $\mathbb{R}$  is apt. A vector space is associated with a field such that there is a product of the members of the field with the vector space members that maps into the vector space. The field is typically looked at (necessarily) as an entity distinct from the elements of the vector space. But if the vector space is  $\mathbb{R}$ , then the nominally extraneous field is isomorphic to the vector space and need not be considered extraneous since the product of a field element and a vector space element can in this case be also thought of as a product between vector space elements, meaning that this particular  $\mathbb{R}$ -vector space is inherently an algebra. The essential feature in recognizing this triviality is the isomorphism between the field and the vector space. Though slightly more complicated, the statement that ‘the union of the frame components of  $\mathbb{R} \oplus \mathbb{R}^3$  (where  $\mathbb{R}^3$  is Euclidean)

is inherently a magma’ is analogous, through the respective frame component isomorphisms with  $\mathbb{R}$  and Euclidean  $\mathbb{R}^3$ . Applying the linearity condition (spacetime should be ‘flat’), the frame itself is then recognized as an algebra.

## The exceptional status of spacetime: are inverse-normed algebras rare?

As we have already observed, considered from the standpoint of its inherent algebraic structure as ultimately expressed by the inverse norm, the magma-derived algebra inherent in Special Relativity has a fundamental similarity to any of the Cayley-Dixon algebras, and indeed to the  $m \times m$  real matrix algebra itself. From this standpoint, Lorentzian geometry is not so strange.

However, based on the criteria specified in Definition 1, it might also be suggested that inverse normed algebras are uncommon. It is true that the Cayley-Dixon algebras all have an inverse norm (given by the Euclidean norm). But although an inverse norm exists on the algebra of real  $m \times m$  matrices (as  $\sqrt{m}[\det(s)]^{1/m}$ ), its subalgebras (i.e., the real finite-dimensional associative algebras) do not in general have inverse norms. For example, for the two dimensional real matrix algebra with the usual matrix product and elements of the form

$$\begin{pmatrix} \alpha & \beta \\ 0 & \alpha \end{pmatrix}, \quad (29)$$

there is no involution that can allow satisfaction of either criterion of Definition 1. While the Spin Factor Jordan algebras have inverse norms (as the Minkowski norm), none of the other formally real Jordan algebras do. Although the algebra built on the real  $m \times m$  matrices where the product is given by the anticommutator multiplied by one-half has an inverse norm (as  $\sqrt{m}[\det(x)]^{1/m}$ ), the Lie algebra built on the real  $m \times m$  matrices (where the product is given by the commutator) does not have a unital hull [14] associated with an inverse norm. This incidentally motivates the question of enumerating all inverse normed algebras. Another task would be the identification of all algebras not satisfying the criteria of Definition 1 but still admitting a solution to the  $\star$ -inverse symmetric functional equation  $f(x^{-\star})f(x) = \|1\|^2$ .

## A comparison of spacetime algebras as regards derivation of the Minkowski norm

A very successful spacetime algebra (STA) is described in [11] (as a “unified mathematical language for physics”), this being the Clifford algebra  $\mathcal{Cl}_{1,3}(\mathbb{R})$ . However, this algebra is chosen because its associated quadratic form is consistent with the Lorentz signature. Thus, it does not have a role in derivation of the Minkowski norm.

The elegant formulation of [13] forwards the planar-spacetime algebra  $\mathcal{Cl}_{2,0}(\mathbb{R})$ , and is demonstrated to be successful in description of various physical processes. A consequence of this choice is the identification of time with a spatial bivector, rather than introducing an extraneous time dimension (such as Minkowski’s *ict*). However, selection of this Clifford algebra is made precisely because the square of the sum of a scalar and a bivector is consistent with a Lorentz signature. Thus, this approach cannot be considered to address the derivation of the Minkowski norm, although it certainly pertains to its interpretation.

The Algebra of Physical Space (APS) [12] is another successful spacetime algebra, this time utilizing  $\mathcal{Cl}_{0,3}(\mathbb{R})$ . It has been suggested that it is particularly natural, since it is the Clifford

algebra arising from Euclidean  $\mathbb{R}^3$ . However, this selection is not unique. That algebra's multivectors, multiplication rules, and further imposed Clifford conjugate, are not already present or uniquely implied by  $\mathbb{R}^3$  itself (and other algebras could as easily be associated with spacetime). Ultimately, it appears that the motivation for introduction of this algebra relates to the fundamental nature of the Pauli matrices, i.e., this choice of algebra is implied by the formalism of quantum mechanics, and in that regard (insofar as it relates to a “derivation” of the Minkowski norm) this approach has a kinship with references cited in the Introduction. On the other hand, it is interesting to note that although APS is not uniquely implied,  $\mathcal{Cl}_{0,3}(\mathbb{R})$  with involution chosen to be the Clifford conjugate satisfies the criteria of Definition 1, and thus has an inverse norm. It is then quickly verified that this inverse norm applied to a paravector yields the Minkowski norm of the paravector.

In contrast to the above Clifford algebras, the magma-derived algebra we forward unambiguously follows from the preexisting linkages between an observer's time and space through their association with  $\mathbb{R}$  and  $\mathbb{R}^3$ . A magma exists as soon as one postulates spacetime as  $\mathbb{R} \oplus \mathbb{R}^3$  with Euclidean  $\mathbb{R}^3$  - it is simply an acknowledgement of the products pre-existing on and between the constituent vector spaces. It is an inherent constituent of the frame, whether or not one chooses to examine the implications of that constituent. In our approach (and that of others [5]), time arises from the co-dimension 1 foliation of spacetime that is an underlying axiom corresponding to the series of observations/predictions that characterize human experience.

### Is Lorentzian spacetime geometry prior to the laws of physics, or does it derive from the laws of physics?

The unexpectedness of the Minkowski norm and its consequences have prompted explanations of its necessity, given that the well grounded physical formalisms in place have a reality potentially independent of the prior choice of spacetime signature [1,2,4-6]. Nevertheless, the particular physical theories employed

for this purpose (classical mechanics, quantum field theory, classical electrodynamics, gravitation) arise from highly specific physical observations (so using them leads to arguments that the observed universe implies the Minkowski norm, rather than an independently derived Minkowski norm implies the observed universe). The argument in [3] is more general, suggesting (among other things) that the Lorentzian signature is required so that the equations underlying successful prediction will be well-posed. However, the latter is not a constructive argument, being more along the lines of *reductio ad absurdum* and evidently based on the anthropic principle. It doesn't tell us where the metric comes from, but only that it must be. In contrast, we have shown that spacetime Lorentzian geometry arises naturally/inevitably from basic ingredients of a frame (since time and space are found to be *physically* joined, why wouldn't that be governed by their inherent *mathematical* linkage?), using a symmetry argument that (compared to the anthropic argument) seems more consistent with the manner in which explanations of physical reality are usually developed and interpreted.

Ultimately, our objective has been to derive Lorentzian spacetime geometry from non-physical principles (other than the axiomatic existence of “time” and “space” with respect to an observer and the concept of “velocity”), so that the equations observed in nature in part *follow* from an independent derivation of the Minkowski norm - rather than the other way around as in the work cited above, not to mention that of Einstein and Poincare. This at least serves to place the spacetime metric on a level co-equal with the cited physical principles.

### Acknowledgments

The author thanks the referees for suggestions that have improved the clarity of the presentation.

### Author Contributions

Conceived and designed the experiments: FG. Performed the experiments: FG. Analyzed the data: FG. Wrote the paper: FG.

### References

- Schutz JW (1989) Affine structure and isotropy imply Minkowski space-time and the orthochronous Poincare group. *Journal of Mathematical Physics* 30 (3):635-638.
- Carlini A, Greensite J (1994) Why is spacetime Lorentzian? *Physical Review D* 49 (2):866-878.
- Tegmark M (1997) On the dimensionality of spacetime. *Classical and Quantum Gravity* 14:L69-L75.
- van Dam H, Ng YJ (2001) Why 3+1 metric rather than 4+0 or 2+2. *Physics Letters B* 520:159-162.
- Itin Y, Hehl FW (2004) Is the Lorentz signature of the metric of spacetime electromagnetic in origin? *Annals of Physics* 312:60-83.
- Yahalom A (2008) The geometrical meaning of time. *Foundations of Physics* 38:489-497.
- Tegmark M (2014) *Our Mathematical Universe*. New York: Alfred A. Knopf.
- Yau S-T (2013) Mathematical object or natural object? In: Brockman J, Editor. *This Explains Everything*. New York:Harper. p. 34.
- Bourbaki N (1974) *Algebra I*. Reading, Mass.: Addison-Wesley. p. 1.
- Page DN, Wootters WK (1983) Evolution without evolution: dynamics described by stationary observables. *Physical Review* 27:2885.
- Hestenes D (2003) *Physics with Geometric Algebra*. *American Journal of Physics* 71:691-714.
- Baylis W, Jones G (1989) The Pauli Algebra approach to Special Relativity. *Journal of Physics A* 22:1-16.
- Chappell J, Iqbal A, Lanella N, Abbott D (2012) Revisiting Special Relativity: a natural algebraic alternative to Minkowski spacetime. *PLOS ONE* 7 (12):1-10 e51756.
- McCrimmon K (2004) *A Taste of Jordan Algebras*. New York: Springer. p. 79.